VIROLOGY 220, 522–529 (1996)
ARTICLE NO. 0342

SHORT COMMUNICATION

Results in Animal-Specific and Rarely Tissue-Specific Selection of *nef* Variants

GE WU ZHU, SAMPA MUKHERJEE, MANISHA SAHNI, OPENDRA NARAYAN, and EDWARD B. STEPHENS[1]

Department of Microbiology, Molecular Genetics, and Immunology, Marion Merell Dow Laboratory for Viral Pathogenesis, University of Kansas Medical Center, Kansas City, Kansas 66160-7240

Received February 13, 1996; accepted April 11, 1996

We analyzed the sequence of *nef* genes from different tissues of three rhesus macaques that had been infected with molecularly cloned $SIV_{mac}239$ for 88 to 92 weeks. Comparison of the predicted amino acid sequences revealed that each macaque had selected out specific amino acid substitutions and that most of this variation (70%) was confined to four regions, amino acids 39 to 75, 90 to 105, 153 to 167, and 191 to 217, comprising 36% of the protein. The *nef* genes in these animals underwent extensive genetic variation with average nucleotide and amino acid substitution rates varying from 0.86 to 2.84% and 2.47 to 6.27%, respectively, although tissue-specific selection of *nef* variants occurred in only 1 of 14 tissues examined in this study. Comparison of the rate of nucleotide and amino acid substitutions in the *nef* genes to those previously reported in the *env* in the central nervous system (CNS) and lymph node (LN) revealed that the predicted amino acid substitution rates for Nef were much higher than for the gp120 region of *env* in the CNS and LN tissues for one macaque. In the two other macaques, the predicted amino acid substitution rates were similar between these two proteins in LN tissues, but the amino acid substitution rates in Nef were significantly higher than in the gp120 from the CNS. Comparison of the nucleotide substitutions in the region of overlap between the *env* and the *nef* revealed that approximately 83% of the nucleotide substitutions in this area resulted in a Nef amino acid sequence change, 26% of the nucleotide substitutions resulted in a gp41 amino acid change, and 9.5% of nucleotide substitutions resulted in amino acid sequence changes in both proteins, suggesting a preference for the selection of amino acid substitutions in the Nef in these animals. Our results indicate that in animals infected with $SIV_{mac}239$ for prolonged periods, variation in the *nef* occurs at rates similar to or exceeding that observed for the *env* gene.

The genomic organization of the primate lentiviruses (HIV-1, HIV-2, and SIV) differs from the ungulate and feline lentiviruses by the presence of a *nef* gene at the 3′ end of the envelope glycoprotein gene (*5, 17*). The *nef* gene of HIV-1 codes for a 27-kDa membrane bound protein that is both myristylated and phosphorylated (*2, 10, 11*). While the *nef* gene has been shown to be dispensable for virus replication in cell culture systems, inoculation of adult macaques with $SIV_{mac}239$ virus containing a deletion in the *nef* gene resulted in reduced virus loads and attenuated virulence for adult macaques (*7, 13*). Recent evidence indicates that Nef down-regulates CD4 expression from the surface virus-infected cells (*4, 9, 10*). Molecular genetic studies have shown that amino acid residues in the cytoplasmic domain of the CD4 molecule are required for Nef-mediated endocytosis of CD4 (*1, 8, 9*). The down-regulation of surface CD4 by Nef appears to act through a common pathway since the endocytosis of murine and human CD4 can be mediated by HIV-1 Nef (*3*). Other studies have shown that Nef enhances virus replication (*16, 24*) and may act at the level of stimulation of proviral synthesis (*22*). More recent studies indicate that Nef is involved in the activation of resting lymphocytes in the host, since the introduction of small numbers of specific mutations in the *nef* gene resulted in a virus that had mitogenic effects in T cells and caused hyperacute disease in macaques (*6*).

Unlike studies on the *env*, only minimal emphasis has been placed on studies of the natural history of the *nef* during prolonged infection in individual hosts. Sequence analyses of the *nef* gene from HIV-1, HIV-2, and SIV isolates have shown that this gene is heterogenous with strong sequence conservation in the four blocks and in the myristylation addition site at the amino terminus, in the region of acidic residues that was postulated to lie on the external surface of the molecule, and a potential recognition site for phosphorylation (R-T-M-S-Y-K) by phosphokinase C (*23*). However, it is not known whether there is a preference for amino acid substitutions in certain regions of the protein or if there is a tissue-specific selection of *nef* variants with prolonged infection. This can only be determined in animals inoculated with molecularly cloned virus in which the nucleotide sequence

[1] To whom correspondence and reprint requests should be addressed. Fax: (913) 588-5599.

0042-6822/96 $18.00

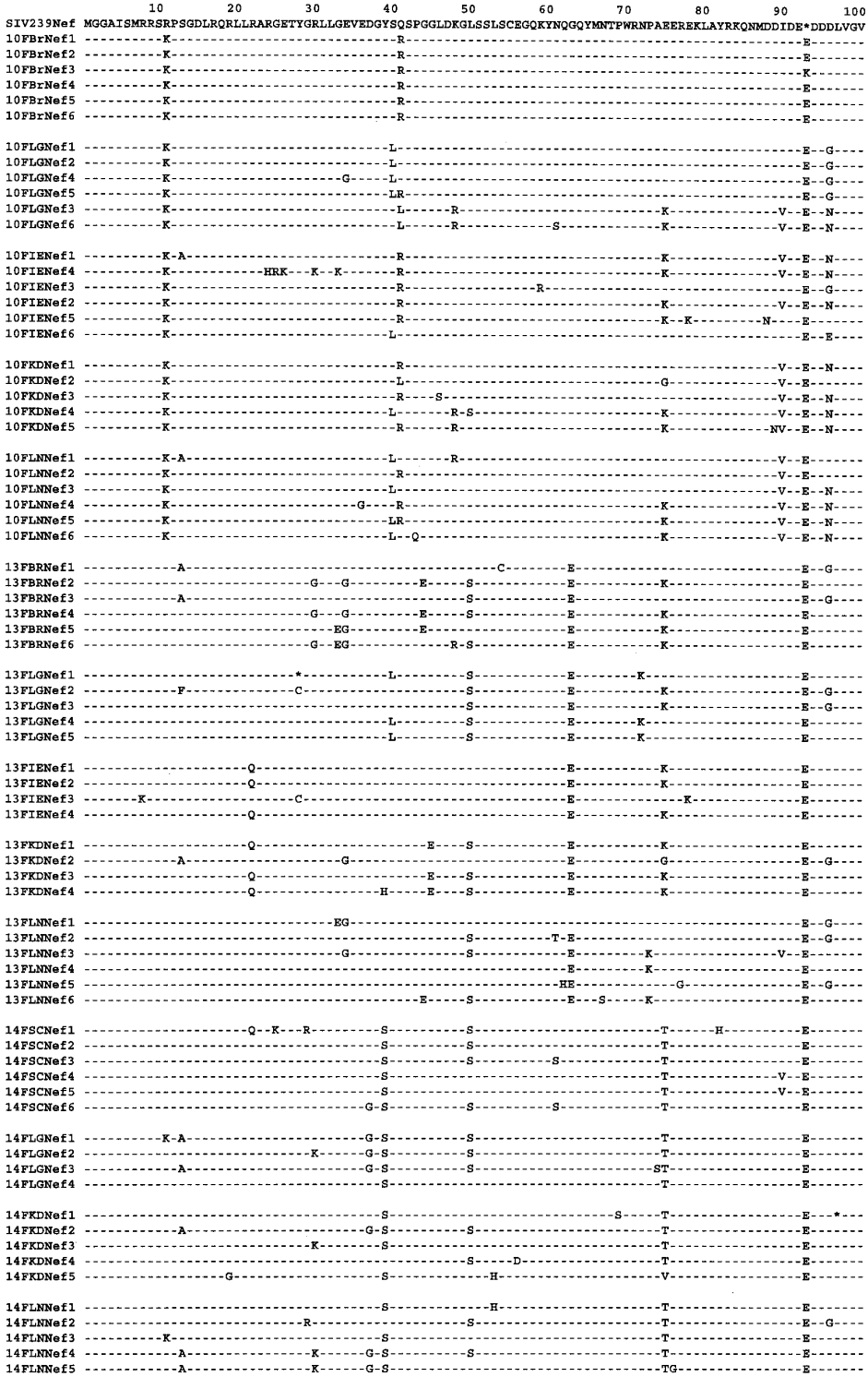

FIG. 1. The predicted amino acid sequence of nef clones isolated from various tissues of macaques 10F, 13F, and 14F at necropsy. Br refers to the frontal cortex, SC to spinal cord, LG to lung, IE to intestine, KD to kidney, and LN to lymph node as the tissue source of DNA for PCR amplification and cloning. Shown above the sequences of the 74 clones is the sequence of $SIV_{mac}239$ Nef protein (*18*). Dashes (-) represent amino acid identity and * represents a stop codon.

of virus is known. Molecularly cloned $SIV_{mac}239$ causes persistent infection and AIDS in most inoculated macaques within 2 years (*14*). In 6- and 10-month studies of macaques infected with SIV_{mac}, only minimal, randomly distributed nucleotide substitutions were observed in *nef* with no tissue-specific changes found (*15, 19*).

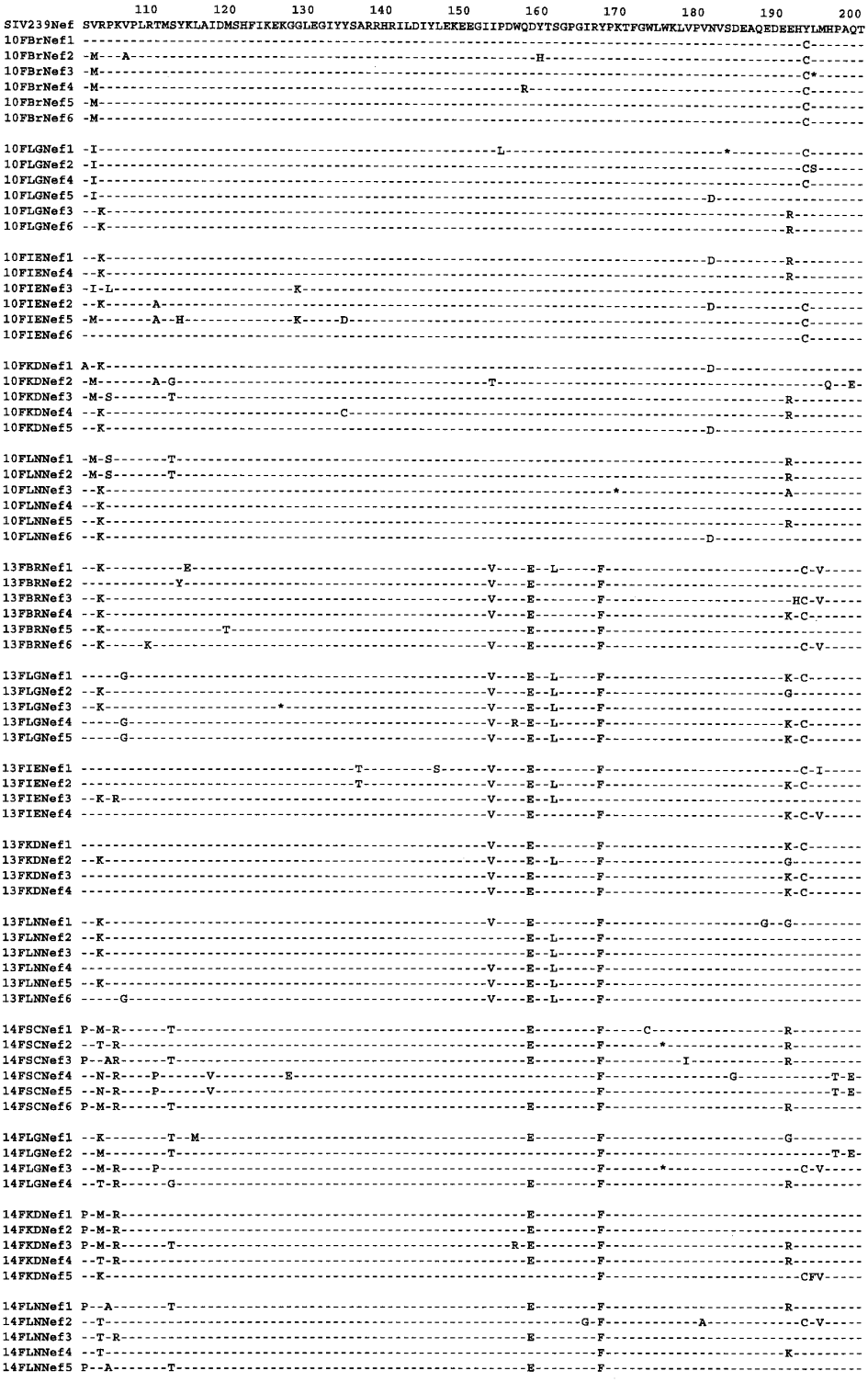

FIG. 1—*Continued*

In this report, we analyzed *nef* gene sequences from three macaques (macaque 10F, 13F, and 14F) that had been inoculated with molecularly cloned $SIV_{mac}239$ and maintained for 89, 88, and 92 weeks, respectively. The pathogenesis of the infection in the animals with respect to tissue tropism and nucleotide substitution rates of the *env* genes from different tissues has been described previously and provided the basis of the present report (*12, 25*). In that study, we showed that the gp120 region of *env* isolated from the CNS of some macaques had a very low nucleotide substitution rate when compared to those isolated from lymph node tissue, whereas in other

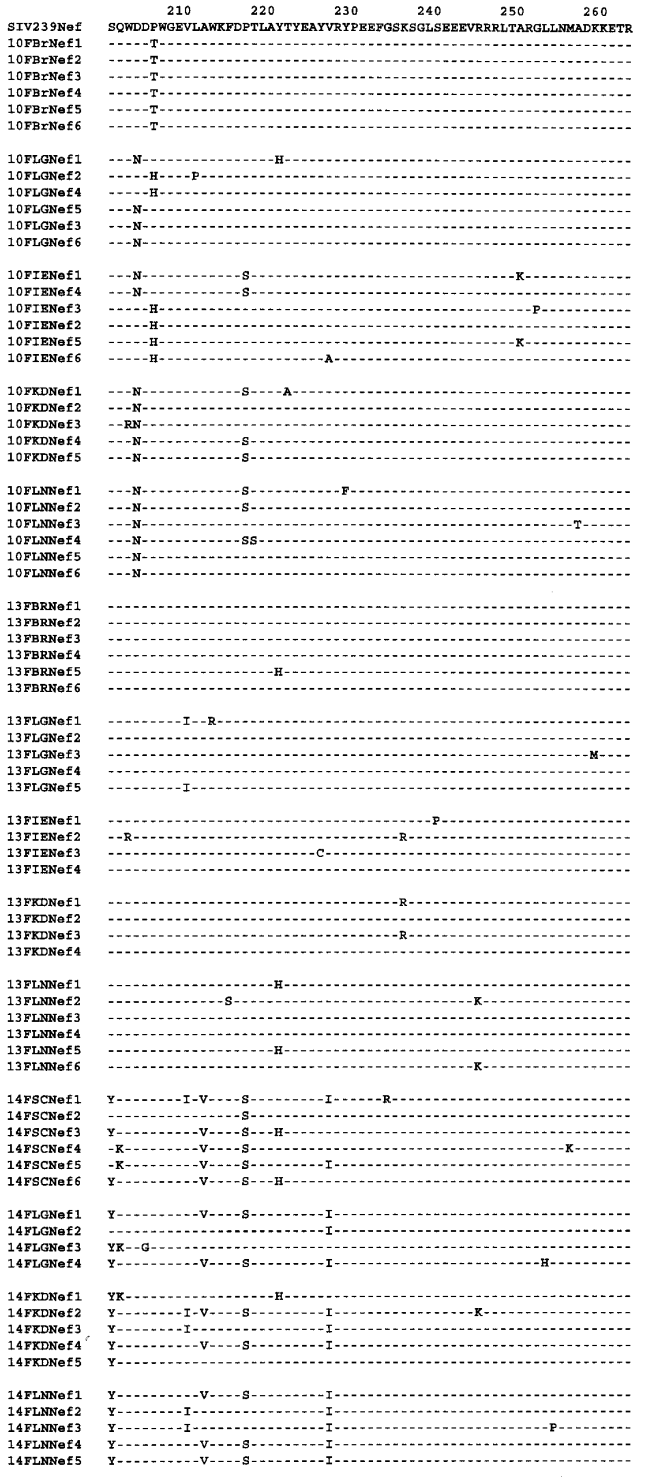

FIG. 1—*Continued*

macaques the nucleotide substitution rates between the two tissues were similar (*25*). In this study, we determined whether the *nef* gene underwent variation at the same rate as *env* in different tissues. Total cellular genomic DNA was extracted from one region of the central nervous system and three to four nonneural tissues (lung, mesenteric lymph node, intestine, and kidney) and used as a template in a nested polymerase chain reactions (PCR) to amplify the SIV_{mac} *nef* gene. The oligonucleotide primers used in the first round were 5′-CAGGACTGA-ACTGACCTACCTACAATATGG-3′ and 5′-ACATCCCCT-TGTGGAAAGTCCCTGCTGTTT-3′ (noncoding), which are complementary to bases 9051 to 9080 and 9868 to 9898 of the SIV_{mac}239 genome, respectively (*18*). One microgram of genomic DNA was used in the PCR (*20, 21*) containing 2.0 m*M* $MgCl_2$, 200 μM each of the four deoxynucleotide triphosphates, 100 pM each oligonucleotide primer, and 2.5 U of Taq polymerase (Perkin–Elmer–Cetus, Norwalk, CT). The template was denatured at 92° for 3 min and PCR amplification performed with an automated DNA Thermal Cycler (Perkin–Elmer–Cetus) for 35 cycles using the following profile: denaturation at 92° for 1 min, annealing at 55° for 1 min, and primer extension at 72° for 3 min. Amplification was completed by incubation of the PCR for 10 min at 72°. One microliter of the PCR product from above was used in a nested PCR using the same reaction conditions described above. For the second round of amplification, the nested set of primers was 5′-ATGGGTGGAGCTATTTCCATGAGGCGGTCC-3′ and 5′-TCAGCGAGTTTCCTTCTTGTCAGCCAT-3′ (noncoding), which are complementary to bases 9077 to 9106 and 9842 to 9868 of the SIV_{mac}239 genome, respectively. The amplified *nef* gene products from three separate PCRs were combined from each tissue and molecularly cloned into the pGEM-T vector (Promega). Plasmids containing *nef* gene sequences were sequenced by primer-directed dideoxy sequencing using Sequenase enzyme (U.S. Biochemicals, Cleveland, OH) and [^{35}S]dATP. SIV_{mac} *nef* sequences were compared to the original SIV_{mac}239 sequence using PCGENE sequence analysis software programs.

The predicted amino acid sequences of the 74 *nef* clones are shown in Fig. 1 and a summary of the genetic changes in the genes isolated from macaques 10F, 13F, and 14F are shown in Tables 1, 2, and 3, respectively. No nucleotide deletions or insertions were observed in any of the *nef* clones (Fig. 1). Only 5 of 74 clones (6.7%) had premature termination codons. The substitution rates of the *nef* clones varied among the three animals; macaque 14F had the highest nucleotide substitution rate (mean 2.45%) followed by macaques 13F and 10F, which had nucleotide substitution rates of 1.69 and 1.53%, respectively (Tables 1, 2, and 3). The average nucleotide and predicted amino acid substitution rates also varied with each tissue of the infected macaques. The average nucleotide substitution rate for macaque 10F ranged from 0.86% for clones derived from the CNS clones to 1.91% for *nef* clones derived from the intestine. For macaque 13F, the average nucleotide substitution rates ranged from 1.45% for *nef* clones derived from the lymph node to 2.1% for clones derived from lung tissue. For macaque 14F, the average nucleotide substitution rates ranged from 2.27% for *nef* clones derived from the

TABLE 1

Genetic Changes in the *nef* Genes Isolated from Macaque 10F Compared to $SIV_{mac}239$

Tissue	No. of clones sequenced	Total No. of nucleotides sequenced	Total No. of nucleotide substitutions	Average nucleotide substitution rate (%)	Average amino acid substitution rate (%)	Nucleotide substitutions leading to an amino acid change (%)
CNS[a]	6	4752	41	0.86	2.47	88.10
Lung	6	4752	68	1.43	3.49	80.80
Intestine	6	4752	91	1.91	4.44	76.92
Kidney	5	3960	70	1.76	4.49	84.29
Lymph node	6	4752	82	1.73	4.06	78.05
Total	29	22,968	352	1.53	3.79	81.63

[a] Viral DNA was amplified from 10F frontal cortex tissue DNA.

lymph node tissue to 2.84% for clones derived from the CNS tissue. The observed variation in *nef* was not the result of $SIV_{mac}239$ replication *in vitro* since transfection of the genome and growth of the virus in macaque PBMC for 1 week resulted in only one nucleotide change in the two clones analyzed (a nucleotide substitution rate of 0.06%; data not shown).

Comparison of the predicted amino acid sequences shown in Fig. 1 indicated that most of the amino acid heterogeneity (70%) was confined to 25 positions in four regions comprising only 36% of the protein sequence. These included six amino acids from residues 39 to 75 (specifically, residues 39, 40, 41, 50, 63, and 75), seven amino acids from residues 90 to 105 (specifically, residues 90, 93, 96, 101, 102, 103, and 105), four amino acids from residues 153 to 167 (specifically, residues 153, 158, 161, and 167), eight amino acids from residues 191 to 217 (specifically, residues 191, 193, 195, 201, 204, 206, 212, and 217). In a previous study that compared the amino acid sequences of HIV-1, HIV-2, and SIV Nef proteins, four blocks of conserved amino acids were identified (*23*). These blocks included amino acid residues from 95 to 129 (block A), 137 to 147 (block B), 164 to 182 (block C), and 209 to 220 (block D) (*23*). Our results indicate that amino acid residues 95 to 114 of block A displayed considerable variability (16.5% of the substitutions but representing only 5.7% of the sequence), whereas residues 115–129 were highly conserved. While blocks B and C (with the exception of a tyrosine to serine substitution at position 167) were well conserved but the conservation of amino acids in block D appeared to be animal-specific, being conserved in macaque 13F but more heterogenous in macaques 10F and 14F. Our results revealed three additional conserved regions (amino acids 1 to 10, 76 to 88, and 228 to 263) in the predicted amino acid sequences from *nef* genes isolated from these three animals. Collectively, amino acids 1 to 10, 76 to 88, 115 to 153, 168 to 180, and 228 to 263 represent 42% of the protein sequence, but had only 5.0% of the observed amino acid substitutions in the *nef* genes sequenced (Figs. 1 and 2).

The results shown in Fig. 1 indicate an animal-specific selection of *nef* genotypes from each of the three macaques. For example, macaque 10F had an arginine to lysine substitution at position 10 in all 29 clones, macaque 13F had a glycine to glutamic acid substitution at position 63 in 24 of 25 clones, and macaque 14F had a tyrosine to serine substitution at position 39 in 18 of 20 clones examined. In addition, there were consensus amino acid substitutions observed among the *nef* clones

TABLE 2

Genetic Changes in the *nef* Genes Isolated from Macaque 13F Compared to $SIV_{mac}239$

Tissue	No. of clones sequenced	Total No. of nucleotides sequenced	Total No. of nucleotide substitutions	Average nucleotide substitution rate (%)	Average amino acid substitution rate (%)	Nucleotide substitutions leading to an amino acid change (%)
CNS[a]	6	4752	74	1.76	4.69	88.10
Lung	5	3960	83	2.10	5.02	79.52
Intestine	4	3168	51	1.61	4.47	92.16
Kidney	4	3168	58	1.83	4.66	87.93
Lymph node	6	4752	69	1.45	4.00	91.30
Total	25	19,800	335	1.69	4.57	87.80

[a] Viral DNA was amplified from 13F frontal cortex tissue DNA.

TABLE 3

Genetic Changes in the *nef* Genes Isolated from Macaque 14F Compared to $SIV_{mac}239$

Tissue	No. of clones sequenced	Total No. of nucleotides sequenced	Total No. of nucleotide substitutions	Average nucleotide substitution rate (%)	Average amino acid substitution rate (%)	Nucleotide substitutions leading to an amino acid change (%)
CNS[a]	6	4752	135	2.84	6.27	84.44
Lung	4	3168	75	2.37	5.70	86.67
Kidney	5	3960	88	2.22	5.25	85.23
Lymph node	5	3960	90	2.27	5.32	85.56
Total	20	15,840	388	2.45	5.64	85.48

[a] Viral DNA was amplified from 14F spinal cord tissue DNA.

from two of three macaques. For example, the leucine to serine substitution at position 50, the isoleucine to glutamic acid substitution at position 153, and the tyrosine to phenylalanine substitution at position 167 was observed in both macaques 13F and 14F. The finding of the same amino acid substitutions in more than one animal suggests a nonrandom occurrence of these substitutions. Of the 14 tissues analyzed from these three animals, tissue-specific selection of *nef* variants were not observed except for those clones isolated from macaque 10F brain. Clones isolated from macaque 10F brain all had a tyrosine to cysteine substitution at position 193 and a proline to threonine substitution at position 206. Alignment of the predicted Nef sequences from macaque

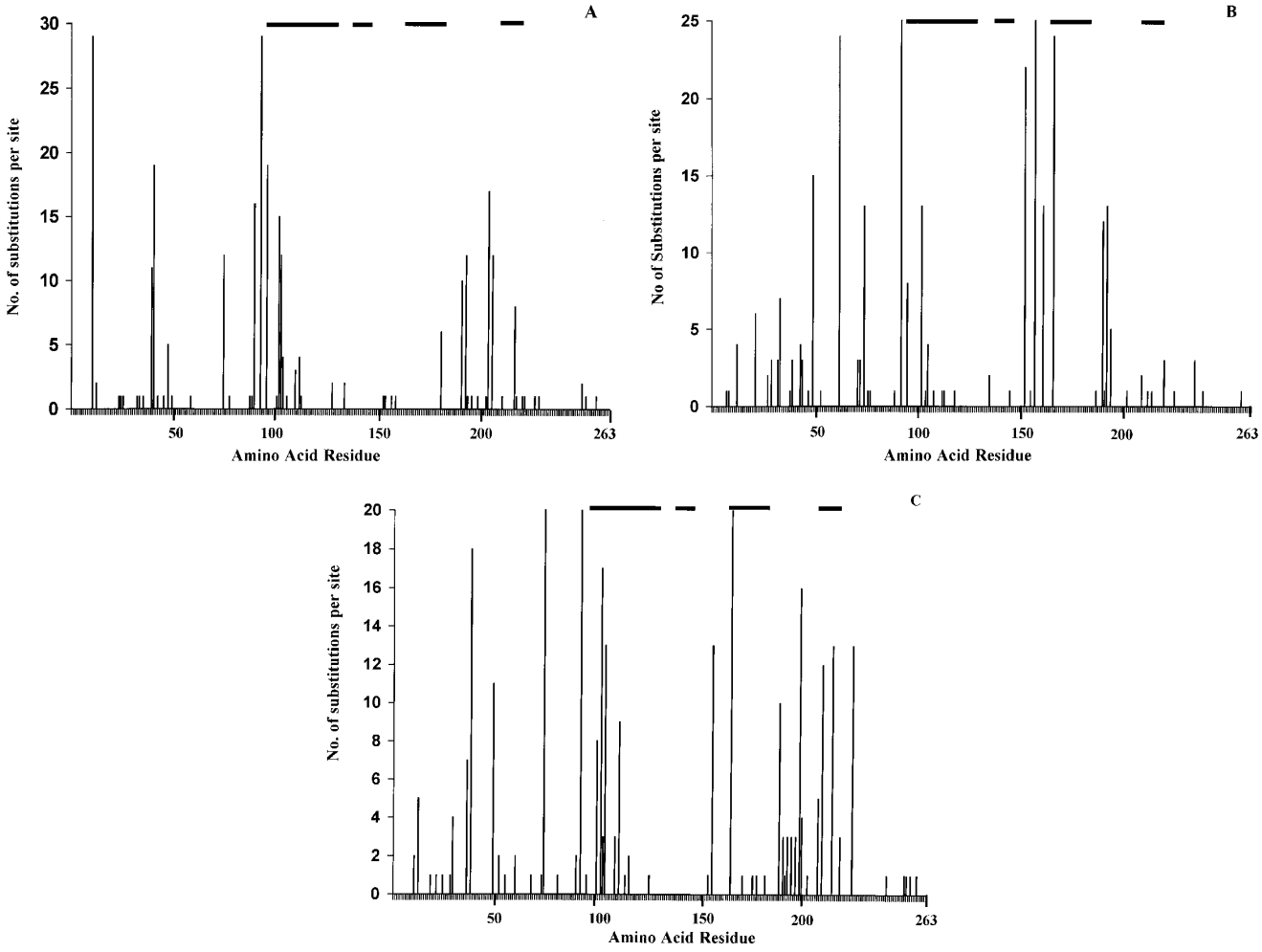

FIG. 2. Schematic diagram of the number of amino acid substitutions at each site for all nef clones isolated from macaques 10F (A), 13F (B), and 14F (C). For comparative purposes, solid bars have been placed over conserved blocks of amino acids previously identified by Shugars *et al.* (*23*).

TABLE 4

Most Frequent Amino Acid Substitutions from the *nef* Genes Isolated from Macaque 10F, 13F, and 14F

Animal	Amino acid position	Amino acid substitution	No. positive/total No. clones sequenced
10F	12	R → K	29/29
	41	Q → R/L	21/29
	90	I → V	16/29
	93	Ter. → E/K	29/29
	96	D → V/G/E	19/29
	102	V → M/I	15/29
	204	D → N	17/29
13F	50	L → S	15/25
	63	G → E	24/25
	75	E → K/G	13/25
	93	Ter. → E	25/25
	103	R → K	13/25
	153	I → V	22/25
	158	D → E	25/25
	161	S → L	14/25
	167	Y → F	24/25
	193	Y → C	13/25
14F	39	Y → S	18/20
	50	L → S	11/20
	75	E → T/V	20/20
	93	Ter. → E	20/20
	103	R → M/T/N/K	17/20
	105	K → R	13/20
	158	D → E	13/20
	167	Y → F	20/20
	191	E → R/K/G	10/20
	201	S → Y	16/20
	212	A → V	12/20
	217	P → S	13/20
	227	V → I	13/20

10F also showed that the sequences isolated from the brain clustered together (data not shown). The most frequent amino acid substitutions occurring in the *nef* genes from each of the three macaques are listed in Table 4.

We compared the nucleotide and amino acid substitution rates sequences of the *nef* gene isolated from CNS and lymphoid tissues with the gp120 genes isolated from the same tissues and published previously (*25*). The results of this comparison are summarized in Table 5. The results indicated that the nucleotide and amino acid substitution rates of the *nef* genes isolated from one animal (14F) were significantly higher than the gp120 genes derived from both brain and the lymph node tissue, while in the other two animals, 10F and 13F, the nucleotide and amino acid substitution rates were slightly higher in *nef* genes than the gp120 genes derived from CNS but were very similar from lymphoid tissues.

We examined the nucleotide substitutions in the region of the *nef* that overlaps with the gp41 region of *env* to determine if their was a preference for the selection of variants with amino acid substitutions in either the *nef* or *env* gene products. The structure of the *nef* gene in $SIV_{mac}239$ overlaps the gp41 region of the *env* gene by 163 bases or 20.6% of the gene (*18*). Analysis showed that a total of 221 nucleotide substitutions (20.6% of the total substitutions) in the 74 *nef* clones occurred in this region of overlap (Table 6). Approximately 82.6% of the 221 nucleotide substitutions led to an amino acid change in *nef* gene product. Correspondingly, only 26.0% of the nucleotide substitutions in the region of overlap with *env* led to an amino acid change in gp41. Only about 9.5% of substitutions in this region led to an amino acid change in both gp41 and Nef. These results indicated a preference for the selection of variants with amino acid substitutions in Nef versus the gp41 in this region.

There are several points that suggest that the *nef* genes analyzed in this study arose from actively replicating virus and were part of an evolutionary process. First, there was a reversion of the premature termination codon at position 93 to a glutamic acid in every *nef* clone analyzed. If the *nef* genes analyzed were from defective viruses, there would not have been selective pressure to revert this termination codon to one coding for an amino acid. Second, if the amplified products were from defective viruses (and hence unable to replicate and spread), one would expect these viruses to be localized in the tissue that they were derived. However, our results indicate that the majority of the observed nucleotide substitutions (and predicted amino acid substitutions) were observed in *nef* genes amplified from multiple tissues. Third, *nef* genes with similar sequence were amplified from multiple PCRs (three used for each tissue), suggesting that they were the predominant *nef* quasispecies in these tissues.

In summary, we have shown that following inoculation of macaques with $SIV_{mac}239$, each animal selected variants with substitutions targeted to specific amino acid residues. Most of the variation (70% of the amino acid changes) was confined to four specific regions of the *nef* gene comprising only 36% of the protein. The rate of

TABLE 5

Comparison of Genetic Changes in the *env* (gp120) and *nef* Genes Isolated from Macaque 10F, 13F, and 14F

		Average nucleotide substitution rate (%)		Average amino acid substitution rate (%)	
Tissue	Animal	*env*	*nef*	gp120	Nef
CNS	10F	0.40	0.86	1.18	2.47
	13F	1.46	1.76	3.70	4.69
	14F	0.51	2.84	0.99	6.27
LN	10F	1.68	1.73	4.48	4.06
	13F	1.19	1.45	4.0	4.0
	14F	1.28	2.27	3.0	5.32

TABLE 6

Analysis of the Nucleotide Substitutions That Overlap the env and *nef* Genes

Animal	No. of nef clones sequenced	Total nucleotide substitutions in overlap region	Rate of nucleotide substitutions leading to Nef a.a. change (%)	Rate of nucleotide substitutions leading to gp41 a.a. change (%)	Rate of nucleotide substitutions leading to an a.a. change in both (%)
10F	29	86	90.70	12.79	4.65
13F	25	67	77.61	34.33	11.94
14F	20	68	79.41	30.88	11.76
Total	74	221	82.57	26.00	9.45

nucleotide and amino acid variation was comparable to and in some cases exceeded the observed variation for the gp120 region of the *env* gene. However, while it appears that tissue-specific selection of virus variants occurs in the *env* gene (*15, 25, 26*), tissue-specific selection of *nef* genotypes was less prevalent.

ACKNOWLEDGMENT

The work reported here is supported by Grants AI-29382, NS-32203, RR-06753, and DK-49516 from the National Institutes of Health.

REFERENCES

1. Aiken, C., Konner, J., Landau, N. R., Lenburg, M. E., and Trono, D., *Cell* 76, 853–864 (1994).
2. Allan, J. S., Coligan, J. E., Lee, T. H., McLane, M. F., Kanki, P. J., Groopman, J. E., and Essex, M., *Science* 230, 810–813 (1985).
3. Anderson, S., Shugars, D. C., Swanstrom, R. C., and Garcia, J. V., *J. Virol.* 67, 4923–4931 (1993).
4. Benson, R. E., Sandridson, A., Ottinger, J. S., Doyle, C., and Cullen, B. R., *J. Exp. Med.* 177, 1561–1566 (1993).
5. Cullen, B. R., *Virology* 205, 1–6 (1994).
6. Du, Z., Lang, S. M., Sasseville, V. G., Lackner, A. A., Ilyinskii, P. O., Daniel, M. D., Jung, J. U., and Desrosiers, R. C., *Cell* 82, 665–674 (1995).
7. Fisher, A. G., Ratner, L., Mitsuya, H., Marselle, L. M., Harper, M. E., Broder, S., Gallo, R. C., and Wong-Staal, F., *Science* 233, 655–659 (1986).
8. Garcia, J. V., Alfano, J., and Miller, A. D., *J. Virol.* 67, 1511–1516 (1993).
9. Garcia, J. V., and Miller, A. D., *Nature* 350, 508–511 (1991).
10. Guy, B., Kieny, M. P., Riviere, Y., LePeuch, C., Dott, K., Girard, M., Montagnier, L., and Lecocq, J. P., *Nature* 330, 266–269 (1987).
11. Harris, M. P. G., and Neil, J. C., *J. Mol. Biol.* 241, 136–142 (1994).
12. Joag, S. V., Stephens, E. B., Adams, R. J., Foresman, L., and Narayan, O., *Virology* 200, 436–446 (1994).
13. Kestler III, H. W., Ringler, D. J., Mori, K., Panicali, D. L., Sehgal, P. K., Daniel, M. D., and Desrosiers, R. C., *Cell* 65, 651–662 (1991).
14. Kestler, H., Kodama, T., Ringler, D., Marthas, M., Pedersen, N., Lackner, A., Regier, D., Sehgal, P., Daniel, M., King, N., and Desrosiers, R. C., *Science* 248, 1109–1112 (1990).
15. Kodama, T., Mori, K., Kawahara, T., Ringler, D. J., and Desrosiers, R. C., *J. Virol.* 67, 6522–6534 (1993).
16. Miller, M. D., Warmerdam, M. T., Gaston, I., Greene, W. C., and Feinberg, M. B., *J. Exp. Med.* 179, 101–113 (1994).
17. Ratner, L., and Niederman, T. M. J., *Curr. Top. Microbiol. Immunol.* 193, 169–208 (1995).
18. Regier, D. A., and Desrosiers, R. C., *AIDS Res. Hum. Retroviruses* 6, 1221–1231 (1990).
19. Rud, E. W., Cranage, M., Yon, J., Quirk, J., Ogilvie, L., Cook, N., Webster, S., Dennis, M., and Clarke, B. E., *J. Gen. Virol.* 75, 529–543 (1994).
20. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N., *Science* 230, 1350–1354 (1985).
21. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. *Science* 239, 487–491 (1988).
22. Schwartz, O., Marechal, V., Danos, O., and Heard, J.-M., *J. Virol.* 69, 4053–4059 (1995).
23. Shugars, D. C., Smith, M. S., Glueck, D. H., Nantermet, P. V., Seiller-Moiseiwitsch, F., and Swanstrom, R., *J. Virol.* 67, 4639–4650 (1993).
24. Spina, C. A., Kwoh, T. J., Chowers, M. Y., Guatelli, J. C., and Richman, D. D., *J. Exp. Med.* 179, 115–123 (1994).
25. Stephens, E. B., Liu, Z. Q., Zhu, G. W., Adany, I., Joag, S. V., Foresman, L., Berman, N. E. J., and Narayan, O., *Virology* 213, 600–614 (1995).
26. Zhu, G. W., Liu, Z. Q., Joag, S. V., Pinson, D. M., Adany, I., Narayan, O., McClure, H. M., and Stephens, E. B., *J. Neurovirol.* 1, 78–91 (1995).